Ben Bailey
Current Issues in Cognitive Science
Mark Feinstein
2015-12-18

An Escalation Model of Consciousness

Introduction

The idea of consciousness has plagued humanity since its inception. Humans as a species have been creating various models to explain this key phenomena of their existence as long as it has been possible. What follows is an attempt to synthesize several popular ideas of cognitive science into a basic conceptual theory of consciousness that represents the state of the field. These ideas include Brooks' subsumption architecture of artificial intelligence and the idea that unique properties of systems can emerge from the interaction of several smaller parts of those systems from complexity sciences. This essay will then examine the implications of this theory of consciousness on psychedelic studies, artificial intelligence, and genetic programming.

The Model

The escalation model of consciousness has four layers:

1. Input layer
2. Perceptual layer

3. Intelligent layer
4. Linguistic/conscious/spiritual layer

In this model, consciousness arises from the interaction of several subsystems or layers of computation/intelligence as a complex phenomenon. It does not have any specific sort of "neural correlate," but instead is created through the *interaction* of these sub-intelligences, or intelligence layers.

At a very basic level, the process goes like this: an organism experiences a stimulus, then the organism's senses decide whether to pass that stimulus on to the intelligence layer or reject it. The intelligence layer is made up of many sub-intelligences which pass the signal on to other sub-intelligences. This interaction creates a conscious layer, which is the organism's conscious experience of life.

With that basic understanding of the model, we can walk through a more specific typical "thought process" in the escalation model. This model is called the *escalation model* due to the ability at each step of the model for the current layer to either *deescalate* or *escalate* a piece of information or an input to the next layer after a modification to that input.

First, the organism receives an input through an "input site," such as the nose or any number of touch receptors located along the human body. Then, that input is either passed along to an intelligence layer or simply rejected if it does not pass a certain threshold by the perceptual layer. The organism-wide intelligence layer is made up of many sub-intelligences that make decisions based in a small problem domain. The sub-intelligences act individually, though some are higher in the chain of sub-intelligences

than others and can influence the outputs of sub-intelligences below them. The combined output of these sub-intelligences is the output from the "intelligent" layer. The interactions of these sub-intelligences as a complex system create the unique phenomenon that we experience as consciousness. This layer has access to the outputs of the intelligence layer—it treats those outputs as "thoughts." In a way, it is simply another sub-intelligence, but it is at the top of the hierarchy and we have an experience of it. The layer that arises from the interactions of the sub-intelligences makes up for key issues with the subsumption architecture, namely the lack of memory and symbolic representation, which is necessary for language. Since many hold the view that linguistic capability is necessary for consciousness, this layer will be referred to as the "linguistic" layer, though it can also be called the "spiritual" layer for reasons we'll see later.

Now we can examine how this model of consciousness could possibly affect different fields of cognitive science in the future, starting with psychedelic studies.

Psychedelic Consciousness

This theory has several implications for the field of psychedelic studies and mystical experiences. The question arises: which layers of the model to classical psychedelics affect?

It has been shown that classical psychedelics increase the basic "connectivity" level of the brain (Petri et al. 2014). One study claims that psilocybin can be used to

help study "neurobiological basis of altered states of consciousness" due to this phenomena (Hasler, Grimberg, Benz, Huber, & Vollenweider, 2004).

So what part of the classical psychedelic experience engenders these altered states of consciousness in the escalation model? I argue that the reason our experience of consciousness changes so greatly during the psychedelic experience is that the linkages between our layers of intelligence are greatly altered. It makes sense that if the connections between layers of intelligence are altered, so will our experience of the phenomenon with those interactions create.

This plays a big role in why the conscious or linguistic layer can also be referred to as the spiritual layer. As psilocybin has been shown to create an experience quite similar to the typical mystical experience. Both experiences rely heavily on altering consciousness, which would suggest that spirituality is extremely dependent on how organisms experience consciousness. Similarly, this layer is very closely tied to how dualists describe the soul, while the intelligence layer is closer to how they describe the mind. Thus, it seems reasonable to describe the conscious layer as spiritual.

Artificial Consciousness

The theory also poses interesting questions about artificial intelligence. Since it is grounded in Brooks' subsumption architecture, the model obviously draws a lot from the field of artificial intelligence. To understand the escalation model of consciousness, we must first understand Brooks' subsumption architecture. Brooks writes that in his robots, "each layer in the subsumption architecture is composed of a fixed-topology

network of simple finite state machines" (Brooks 1991), with each layer being combined through suppression or inhibition. Suppression and inhibition both involve "side-tapping" the signal from the input layer, and either suppressing that layer's behavior for a short period of time and passing on a new message, or simply inhibiting the layer's output.

In the case of humans and other biological organisms, each layer is not exactly a network of finite state machines, but rather a collection of neurons and cells which group together to form a specific purpose. Some strict cognitive scientists may argue that those *are* finite state machines, but the exact definition of a sub-layer of consciousness is not necessary to continue the argument.

The implications of this theory for artificial intelligence are fairly large. The theory would suggest that a robot could certainly become conscious if the interactions between its sub-layers of intelligence were complex enough. This creates a slew of ethical issues that we as a species are most likely not ready to deal with yet. But how do simple interactions create complex phenomena?

Complex Consciousness

This theory is grounded not only in Brooks' subsumption architecture and other artificial intelligence research, but also complexity sciences. It takes a large part of its definition of consciousness from Tononi's Integrated Information Theory, which defines the quantity of consciousness as corresponding to "the amount of integrated information generated by a complex of elements," integrated information being "the amount of

information generated by a complex of elements, above and beyond the information generated by its parts" (Tononi 2008).

Tononi's definition of consciousness draws heavily from the rising field of complexity sciences. The idea that complex systems can engender properties that are greater than the sum of their parts is a revolutionary one.

We can combine Brooks' revolutionary subsumption architecture with Tononi's Integrated Information Theory to come up with a definition of *human* consciousness as such: "the amount of information generated by the multitude of layers of intelligence that cannot be generated by those layers alone."

It is important to specify "human" consciousness here, as it is difficult for us as humans to try and define consciousness outside of how we experience it. Not only are we presented with the challenge of defining something that cannot be defined without itself, but we then have to attempt to extrapolate that across different forms of life and, possibly, even ordinary objects.

Evolving Consciousness

Finally, we can examine the implications of the model on the field of evolutionary programming. Evolutionary programming is a type of programming where the engineer creates a program that has the capacity to evolve itself. This enables evolution to create unique and interesting solutions to problems that humans could not figure out. A good example of this could be search algorithms. We've spent a long time trying to perfect search algorithms over the last few decades, and while we've gotten pretty good at

efficient searching, it's definitely possible that there is a faster method that relies on ways of thinking that don't come naturally to humans.

One of Brooks' major ideas was that programming a sort of "central representation" of a world was a red herring in artificial intelligence programming. Part of his reasoning for this is that manually creating a central representation of a robot's environment would either be a gargantuan task or the robot would be limited to a very small, specific environment such as the "blocks" world.

Brooks proposed that his subsumption architecture could create a robot that behaved similarly to a robot that had a representation of the world programmed inside of it. This assumption proved mostly correct, and one can observe a similar phenomenon with wolf hunting behavior. Simple rules governing actions can create behavior that looks very similar to that which could be a result of central planning and a central representation of the world inside wolves.

I propose that, similar to Tononi's Integrated Information Theory, every being with sub-intelligences has the possibility for a consciousness to arise from the interactions of those sub-intelligences. In this case, wolves most likely possess a much smaller "level" of consciousness, perhaps even as low as we wouldn't define it as consciousness but rather intelligence. While it may not be semantically valuable to define wolves as conscious, I do believe that it is technically correct, and that we draw an imaginary line between intelligence and consciousness.

What if we evolved consciousness in robots through evolutionary computing? For the purpose of this paper, we will ignore the possibly drastic philosophical and political ramifications of this idea and focus on the practicality.

Once consciousness is simplified down to an emergent property of interacting layers of intelligence, it becomes much simpler to think about how one would go about trying to evolve consciousness in a robot. One could focus on creating layers of intelligence that evolve based on inputs and outputs and see if eventually a form of consciousness emerges from the connections of the evolved layers. It would be even more interesting and compelling to evolve the linkages between the layers as well. The final step could be to evolve the evolutionary process itself, including selection and mutation techniques.

Conclusion

Now that we have an understanding of each individual piece of the model and how it would affect the fields it draws from, we can finally take a specific example through this architecture.

First, let's say a fly lands on an organism's arm. This is the input. The touch receptor in that organism's arm where the fly landed passes the signal to the lowest layer of intelligence that is listening for inputs from that receptor, as the input signal was higher than the designated threshold needed to pass the signal on. This is the perceptual layer. For the sake of this example, we'll make the simplifying assumption that the first sub-intelligence connected to this touch receptor is one which has the goal

to avoid direct unplanned contact. This sub-intelligence passes the "command" to flinch away from the contact both to the muscle sub-intelligence controlling the arm but also to the higher consciousness layer. This creates an awareness of the fly landing and the flinching in the organism's experienced mind, while simultaneously causing the organism to flinch.

Now what would happen if the action of flinching caused the organism to hit its hand rather forcefully on a nearby wall? This would cause another input to a touch receptor in the hand, which would then be passed on eventually to a sub-intelligence which receives pain, then to one which controls pain reactions. This sub-intelligence passes its reaction on to the conscious layer, which then decides to allow the pain reaction and passes it back down the layer hierarchy while simultaneously storing the event in the organism's memory. In the future, the organism may adjust the strength of its flinch reaction based on its proximity to a hard object.

This example demonstrates just how essential the conscious layer is. Without it, the organism would have no ability to learn, and what happens to it in the present would not affect its reactions in the future.

It is obviously difficult to create a definitive definition of consciousness. Some would even argue that it is pointless. I believe that attempting to define that which in essence *is* our very experience is, at the very least, a valuable exercise. While we may never come to a consensus or even have any hard evidence of any definition, thinking about the way we think and experience being can greatly influence many areas of research as described above. At a bare minimum, it can provide interesting new ideas for

creating artificial intelligence and self-evolving programs, which can, in turn, influence the ways we think about ourselves, creating a sort of cognitive science positive feedback loop.¶

Works Cited

Brooks, Rodney A. "Intelligence without Representation." Artificial Intelligence 47, no. 1 (1991): 139–59.

Hasler, Felix, Ulrike Grimberg, Marco A. Benz, Theo Huber, and Franz X. Vollenweider. "Acute Psychological and Physiological Effects of Psilocybin in Healthy Humans: A Double-Blind, Placebo-Controlled Dose–effect Study." Psychopharmacology 172, no. 2 (2004): 145–56.

Petri, G., P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P. J. Hellyer, and Francesco Vaccarino. "Homological Scaffolds of Brain Functional Networks." Journal of The Royal Society Interface 11, no. 101 (2014): 20140873.

Tononi, Giulio. "Consciousness as Integrated Information: A Provisional Manifesto." Biological Bulletin 215, no. 3 (2008): 216–42. doi:10.2307/25470707.